BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Office de la propriété intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An Agency of Industry Canada

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,246,623,** on October 7, 1998, by **McGILL UNIVERSITY,** assignee of Jerry Pelletier and Manjula Das, for "Oligonucleotide Primers that Destabilize Non-Specific Duplex Formation and Reduce Mispriming During cDNA Library Construction and Uses Thereof".

Agent certificateur/Certifying Officer

October 3, 2003

Date

Canada

OPIC CIPO

ABSTRACT OF THE DISCLOSURE

The present invention relates to genetic engineering. More specifically, a method is presented for reducing mispriming during DNA synthesis. In particular, the present invention relates to universal primers which reduce mispriming during cDNA library construction, thereby increasing the proportion of cDNA clones having been primed from the *bona fide* 3' poly A tail. The present invention further relates to the use of the discriminating oligonucleotides of the present invention in other methods such as mRNA purification, PCR-based detection methods and sequencing.

TITLE OF THE INVENTION

OLIGONUCLEOTIDE PRIMERS THAT DESTABILIZE NON-SPECIFIC DUPLEX FORMATION AND REDUCE MISPRIMING DURING cDNA LIBRARY CONSTRUCTION AND USES THEREOF.

FIELD OF THE INVENTION

The present invention relates to genetic engineering. More specifically, a method is presented for reducing mispriming during DNA synthesis. In particular, the present invention relates to universal primers which reduce mispriming during cDNA library construction, thereby increasing the proportion of cDNA clones having been primed from the *bona fide* 3' poly A tail. The present invention further relates to the use of the discriminating oligonucleotides of the present invention in other methods such as mRNA purification, PCR-based detection methods and sequencing.

BACKGROUND OF THE INVENTION

The isolation and rapid mapping of complementary DNAs (cDNAs) is central to characterizing the information that is of significant biological relevance in the genome of an organism. A full length cDNA allows one to predict transcription initiation start sites, translation initation start sites, deduce certain protein characteristics based on primary amino acid sequence, predict transcription termination sites, and visually inspect the 5' and 3' untranslated regions for elements which may be involved in post-transcriptional regulation of gene expression. The analysis of several complete cDNAs of a given gene enables one to gather information on alternative splicing, alternative

promoter usage, and alternative polyadenylation signals - all events known to be important in gene expression regulation. In addition, the comparison of genomic and cDNA sequence is essential to determine exon-intron structure and document the occurrence of RNA editing - a post-transcriptional regulatory mechanism on which there is little information.

The cloning of mRNA into cDNA for the purposes of functional studies is a complex, interrelated series of enzyme-catalyzed reactions involving the *in vitro* synthesis of a DNA copy of mRNA, its subsequent conversion to duplex cDNA, and insertion into an appropriate prokaryotic vector. The procedure involves the following series of steps (outlined in Fig.1):

1) Isolation of high quality mRNA from the tissue or cell line of interest.

2) Annealing of a DNA oligonucleotide, either a mixture of oligonucleotides of random sequence or an oligo d(T) primer, to the mRNA. When full-length cDNAs are required, oligo d(T) is utilized, since this is expected to anneal to the 3' poly (A) tail of the mRNA.

3) Reverse transcriptase is then utilized to prime from the DNA primer and copy the RNA template into cDNA.

4) Second strand synthesis is performed utilizing RNAse H, DNA polymerase I, and DNA ligase.

5) The ends of the cDNAs are polished, prepared for cloning, and the cDNAs are introduced into an appropriate cloning vector.

Although a number of different approaches can be used to generate cDNA libraries, they all suffer from two major problems, often making the isolation of a complete cDNA an arduous task. The cloning of

incomplete cDNAs is widespread, resulting in only partial characterization of mRNA transcripts and significantly increasing the cost and amount of work required to obtain a full-length copy of the cDNA of interest. One major reason why many clones in current cDNA libraries are not full-length is due to mispriming of the oligo d(T) primer (de Fatima Bonaldo et al., 1996, *Genome Res.* 6:791-806). Many eukaryotic mRNAs contain regions of A-rich stretches within their sequence. Thus oligo d(T) primers can anneal to these internal A-rich stretches, and when reverse transcriptase primes from these internal sites, sequence information from the 3' end of the mRNA is lost during the cDNA cloning process (Fig. 1). Although the genetic code of most organisms is composed of ~ 50% guanosine + cytosine residues and 50% of adenosine + thymidine residues, there are well known examples of organisms whose genetic code deviates from this ratio. For example, the genome of the parasite responsible for malaria transmission, *Plasmodium falciparum*, has a genome of >80% adenosine + thymidine residues (Weber, J.L., 1987, *Gene* 52:103-109). This implies that cDNA libraries derived from this organism will contain many truncated, less-than-full-length clones, due to mispriming of the oligo d(T) primer during first strand synthesis. Mispriming is thus a serious hindrance to gene discovery and characterization.

These technical limitations imply that a set of products of variable length are often generated during first strand synthesis. Consequently, a number of truncated clones are present in any given library. Given these cloning complications, interpretations about gene structure are sometimes misleading and cDNA cloning is often inefficient,

costly, and time-consuming - often requiring the sampling of several different libraries.

The actual procedure for generating cDNA libraries has not extensively deviated from the original method of Gubler et al., 1983, *Gene* 25:263-269. A major limitation of the current technology is that a set of products of variable length are often generated during first strand synthesis. Consequently, a number of truncated clones will be present in libraries for any given gene. Priming from the poly (A) tract of mRNAs with oligo d(T) is necessary to obtain a copy of the entire 3' untranslated region. However, it is the experience of many laboratories screening cDNA libraries, that a significant proportion of clones do not have a *bona fide* 3' end, due to misannealing of the oligo d(T) primer to internal A-rich sites. Indeed, cDNAs with 3' truncations are estimated to occur at frequencies of 10-15% in some libraires (de Fatima Bonaldo et al., 1996, *supra*). Such clones are easily recognized by the absence of a *bona fide* polyadenylation signal sequence ~20 nucleotides upstream of the oligo (dA) tail of the cDNA. If enhanced discrimination could be achieved between annealing to the *bona fide* poly (A) tail versus internal A-rich sequences by the Reverse Transcriptase primer, then the frequency of this "mispriming artifact" would be significantly reduced.

Nucleic acid hybridization, in which a DNA or RNA strand binds to its complement to form a duplex structure is a fundamental process in molecular biology. A critical aspect of this process is the specificity of molecular recognition of one strand by the other. Sequence differences as subtle as a single base change are sufficient to enable discrimination of short (e.g. - 14 mer) oligomers, and are frequently used to detect point mutations in genes (Conner et al.,

1983, *Proc. Natl. Acad. Sci. USA.* 80:278-282.). Molecular discrimination of single point changes using oligonucleotides has been well documented and the underlying thermodynamics well characterized (Ikuta et al., 1987, *Nucl. Acids Res.* 15:797-811; Doktycz et al., 1995, *J. Biol. Chem.* 270:8439-8445; Southern et al., 1994, *Nucl. Acids Res.* 22:1368-1373; Saiki et al., 1989, *Proc. Natl. Acad. Sci. USA* 86:6230-6234). However, in many cases, the stability difference between a perfectly matched complement (e.g. - between a poly (A) tail and oligo $d(T)_{15}$) and a complement mismatched at only one base (e.g. - between AAAAAAATAAAAAAA and oligo $d(T)_{15}$) can be quite small, corresponding to as little as 0.5°C difference in their duplex melting temperature (T_ms) (Fig. 2). The longer the oligomer of interest (e.g. an oligo $d(T)_{20}$ primer versus and oligo $d(T)_{15}$ primer) the smaller the effect of a single-base mismatch on overall duplex stability. This limitation in hybridization is the major reason why oligo d(T) primers often hybridize to internal A-rich sequences on mRNA templates during cDNA library construction, and consequently why a large number of clones in such libraries do not contain the *bona fide* 3' end.

Guo et al. (1997, *Nature Biotech.* 15:331-335) have recently shown that increased discrimination of single nucleotide mismatches by oligonucleotides can be achieved by introducing artifical mismatches into the probe oligonucleotide using the base analog 3-nitropyrrole. This base analog acts as a universal nucleoside that hydrogen bonds minimally with all four bases without steric disruption of the DNA duplex (Nichols et al., 1994, *Nature* 369:492-493). Since hydrogen bonding between bases of two complementary strands of DNA are the major thermodynamic forces responsible for maintaining the

integrity of a double strand DNA duplex, base substitutions with analogs with lessened hydrogen bond capacity can function as universal nucleosides (Nichols et al., 1994, *supra*). A number of different nucleoside analogs have been developed which function in this fashion (Millican et al., 1984, *Nucl. Acids Res.* 12:7435-7453; Inone et al., 1985, *Nucl. Acids Res.* 13:7119-7128; Fukada et al., 1986, *Naturforsch. B.* 41:1571-1579; Seela et al., 1986, *Nucl. Acids Res.* 14:1825-1844; Eritja et al., 1986, *Nucl. Acids Res.* 14:8135-8153; Habener et al., 1988, *Proc. Natl. Acad. Sci. USA* 85:1735-1739; Lin et al., 1989, *Nucl. Acids Res.* 17:10373-10383; Francois et al., 1990, *Tetrahedron Lett.* 31:6347-6350; Brown et al., 1991, *Carbohydrate Res.* 216:129-139.). Guo et al. (1997, *supra)* have shown that the introduction of universal analogues into heteropolymeric oligonucleotides during their synthesis, increases the thermal stability (ΔT_m) of hybrids formed between an oligonucleotide with the universal nucleoside and with normal and single-nucleotide variant DNA targets by as much as 200%, as compared to hybrids formed between a wild-type oligonucleotide and normal or single-nucleotide variant DNA targets.

U.S. patent 5,438,131 of Bergstrom et al. teaches oligonucleotides of at least 10 nucleosides, composed of at least two different bases, and containing at least one universal nucleoside and the use thereof to reduce the element of risk and enhance success in screening DNA libraries. The universal base is defined in U.S. 5,438,131 as being a modified nucleic acid base that can base-paired with its ally, one of the common bases A, T, C and G (as well as U). The aim of the universal base is to reduce degeneracy while still preserving the uniqueness of the probe. A variety of compounds have been investigated

as universal bases and a number of them are described in U.S. 5,438,131. In a preferred embodiment, U.S. 5,438,131 relates to oligonucleotides containing universal nucleosides at degenerate positions, such that the oligomer allows bonding to unknown bases, enabling the formation of duplexes with ambiguous or unknown nucleic acid sequences. In a particularly preferred embodiment, U.S. 5,438,131 relates to 3-nitropyrrole nucleoside as the universal nucleoside. U.S. 5,438,131 thus relates to the use of universal nucleosides in order to stabilize duplex formation between heteropolymers of oligonucleotides and a target nucleic acid.

In view of the technical limitations of current methods of cDNA synthesis, there remains a need to destabilize artefactual duplex formation to increase the discrimination between specific and non-specific duplexes, to provide the means to reduce mismatches in general, and more particularly to reduce mispriming during DNA synthesis, cDNA library construction, and PCR applications. The present invention seeks to meet these and other needs.

The present description refers to a number of documents, the content of which is herein incorporated by reference.

SUMMARY OF THE INVENTION

The invention concerns the identification of primer modifications that can destabilize artifactual duplex formation and decrease the number of mismatches between same and its target sequence.

The invention further concerns the identification of primer modifications that improve the binding thereof to a homopolymeric

target sequence. The invention therefore provides oligonucleotides which are better at discriminating between their homopolymeric complementary sequence and a related target sequence. In addition, the present invention provides assays which can be used (and adapted) to identify oligonucleotide modifications that destabilize mismatches.

The invention also concerns the development of primers which decrease mispriming events encountered during DNA synthesis. More specifically, the invention concerns the development of universal primers which decrease the number of internal mispriming events during cDNA generation, improving the efficiency of correct priming from the *bona fide* 3' poly (A) tail.

The present invention further relates to universal primers which reduce the proportion of mismatches during genetic engineering methods such as, for example, mRNA purification, 3' RACE, 5' RACE, PCR, sequencing and the like. In a particularly preferred embodiment, the present invention relates to the incorporation of at least one universal base in a homopolymeric oligonucleotide in order to reduce mismatches to its homopolymeric target sequence. The invention concerns more particularly homopolymeric oligonucleotide primers having modifications which improve their binding to their target sequence.

More specifically, the present invention relates to primers incorporating 3-nitropyrrole modifications.

The invention also concerns assays to identify modifications in oligonucleotides which reduce the proportion of mismatches and mispriming events, comprising a random or rational design of modifications of a chosen primer, a hybridization thereof with its target sequence to form a duplex, a synthesis of DNA priming from this

duplex and an analysis of the synthesized DNA to assess for the presence of mispriming events, wherein the number of mispriming events is compared between the DNA synthesis initiated from modified and control oligonucleotides.

In accordance with the present invention, there is therefore provided a method for destabilizing non-specific duplex formation between an oligonucleotide and a target nucleic acid, wherein at least one of the oligonucleotide and target nucleic acid is a homopolymeric sequence, comprising an incubation of the target nucleic acid with a modified oligonucleotide, wherein the modified oligonucleotide includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

In accordance with the present invention, there is also provided a method for increasing the proportion of full length cDNA clones in a library, comprising a use of a modified oligo d(T) primer during first strand synthesis, wherein the modification decreases or abrogates hydrogen bonding between the modified oligo d(T) primer and a non-specific target sequence, thereby increasing the proportion of full length cDNA clones.

In accordance with another aspect of the present invention, there is provided a method for reducing mispriming events during DNA synthesis comprising a use of a modified oligonucleotide to prime the DNA synthesis, wherein the modification decreases or abrogates hydrogen bonding between the modified primer and a non-specific target sequence, thereby reducing mispriming events.

In accordance with yet another aspect of the present invention, there is provided modified oligonucleotide primers that

destabilize non-specific duplex formation and reduce mispriming during DNA synthesis.

While the method of the instant invention is demonstrated during first strand cDNA synthesis to improve the quality of cDNA libraries by reducing the number of clones containing aberrant 3' ends due to oligo d(T) mispairing, and more specifically using the eIF-4GII mRNA template, the present invention, which has broad utility, is not so limited. Although 3' mispairing is a general problem encountered when generating cDNA libraries from a number of organisms, this problem can be particularly exacerbated when generating cDNA libraries from organisms that have A rich genomes, since the number of internal A-rich stretches will be higher in genes from these organisms. This type of incomplete A-tract is expected to misanneal to oligo d(T) and produce truncated cDNAs during library construction. An example of such an organism is *Plasmodium falciparum*, the parasite responsible for malaria transmission by mosquitos.

In addition, while the instant invention is demonstrated using an oligo d(T)•Z primer (an oligo d(T) primer in which two of the thymine bases are substituted by 3-nitropyrrole), the instant invention is not so limited. Indeed, the position of the modified bases within the examplified oligonucleotide, oligo d(T) primer, can be altered (Fig. 2C) without changing the discrimination between primer and either complementary template or partially complementary template. Indeed, Guo et al. (1997, *supra*) have changed the position of 2 universal nucleosides within a given heteropolymeric oligonucleotide and shown that in many cases increased discrimination between perfect matched template and mismatched template is maintained. Thus, the instant

invention extends to any homopolymeric oligonucleotide (or any oligonucleotide designed to bind to a homopolymeric target sequence) such as an oligo d(T) primer containing modified nucleosides at any position, provided that such modification maintains the discriminating ability of the oligonucleotide. It should be clear to the person of ordinary skill that the present invention further provides the means to assess whether the modifications alter this discriminating activity of the oligonucleotide. It should also be clear that any type of homopolymeric-complementary sequence duplex formation could be improved by the instant invention. In a broad sense therefore, the present invention provides the means and method to improve homopolymeric complementarity sequence duplex formation.

It should be clear to a person of ordinary skill that the present invention has broad implications since it demonstrates that a modification which results in destabilization of a duplex (examplified with oligo d(T), having 2 substitutions, and its poly A target sequence), significantly decreases the proportion of mismatches and of mispriming events. Hence, it is expected that other types of destabilization of the hydrogen bonds between an oligonucleotide and its target sequence would have the same effect. Non-limiting examples of modifications of the oligonucleotide which would result in such a destabilization of the duplex formation, include modifications which reduce or abrogate hydrogen bonding. Non-limiting more specific examples include known base modifications, base analogs (e.g. inosine), universal bases, and partial mismatches. Of course, it will be understood that such modifications should not favor duplex formation with a non-desired target sequence.

It should also be understood that the different modifications of the oligonucleotides encompassed by the present invention can be adapted by the person of ordinary skill to suit particular utilities (i.e. mRNA purification, sequencing).

The present invention should not be limited to the modifications of oligonucleotides with 3-nitropyrrole, since other universal bases are well known in the art. Indeed, in addition to 3-nitropyrrole, a number of universal nucleosides have been synthesized and characterized (Millican et al. 1984, *supra*; Inone et al., 1985, *supra*; Fukada et al., 1986, *supra*; Seela et al.,1986, *supra*; Eritja et al., 1986, *supra;* Habener et al., 1988, *supra*; Lin et al., 1989, *supra*; Francois et al., 1990, *supra;* Brown et al., 1991, *supra*). Thus, the present invention covers any oligo d(T) primer containing any universal nucleoside which allows for enhanced discrimination when hybridizing to perfect versus mismatched templates.

A non-limiting example of an alternative use of this technology is in mRNA purification, by replacing oligo d(T) affinity matrixes currently employed with modified oligo d(T) according to the instant invention. An oligo d(T)•Z affinity matrix would perform the same task, except that binding to internal A-rich stretches would be minimized and could result in a purification method with a higher stringency than currently employed. This matrix, for example, could provide a better selection between eukaryotic mRNA and contaminating mycoplasmic RNA (which is A-T rich). Since mycoplasms often contaminate tissue culture cell lines, co-purification of mycoplasma RNA with eukaryotic mRNA on oligo d(T) column can produce cDNA libraires contaminated with mycoplasma clones.

Often the sequence of a particular RNA must be interogated. Reverse transcriptase (RT), in combination with PCR, can be used to amplify a given region on an RNA template. The use of oligo d(T)•Z as primer in the RT reaction would ensure that the 3' end of the mRNA is represented on the cDNA template. Thus, the present invention can also be incorporated into current 3' RACE (Rapid Amplification of cDNA Ends) protocols, designed to obtain the 3' end of a given clone.

In some cloning protocols, first strand synthesis is followed by homopolymeric tailing of the products utilizing terminal deoxynucleotidyl transferase. For example, dGTP can be utilized to add a homopolymeric stretch of G's at the 5' end of the cDNA. Thus the DNA polymerase utilized in second strand synthesis can take advantage of this G-stretch by priming from an oligo d(C) primer annealed to the G-stretch positioned at the 5' end. This procedure has the advantage of maintaining the sequence at the 5' terminal end of the cDNA, and is also used in 5' RACE strategies to identify the 5' end of mRNAs (Frohman et al., 1988, *Proc. Natl. Acad. Sci. USA* 85:8998-9002; Loh et al., 1989, *Science* 243:217-220) (Fig. 5). One drawback of this approach however is that, since 5' untranslated regions of mRNAs are usually GC rich, the oligo d(C) can misprime from internal G-rich regions, producing less than full-length cDNAs. It is envisaged that the incorporation of universal nucleosides into such homopolymeric primers would be useful to increase the specificity of binding and generating cDNAs which terminate at the bona fide 5' end. Thus, the present invention further relates to cloning procedures or RACE protocols involving priming of second strand synthesis from a homopolymeric tail.

It may be desirable in some PCR protocols to utilize homopolymeric oligonucleotides containing universal nucleosides (or other non-specific duplex destabilizing modifications), to achieve increased discrimination between a target site (or several target sites) of interest when generating a specific product or a set of products (for example use of an oligo d(T) primer to prime DNA synthesis from the A-rich stretch of Alu repeats in humans). Since these products can be developed to be used as genetic markers (by identifying polymorphisms residing with the sequence of the product), changing the specificity of targeting by altering the specificity of the oligo d(T) primer, could result in a more consistent representation of the final PCR products. The present invention thus further relates to the use of universal oligonucleotides or other modified oligonucleotides, during PCR amplification.

DEFINITIONS

Nucleotide sequences are presented herein by single strand, in the 5' to 3' direction, from left to right, using the one letter nucleotide symbols as commonly used in the art and in accordance with the recommendations of the IUPAC-IUB Biochemical Nomenclature Commission.

Unless defined otherwise, the scientific and technological terms and nomenclature used herein have the same meaning as commonly understood by a person of ordinary skill to which this invention pertains. Generally, the procedures for cell cultures, infection, molecular biology methods and the like are common methods used in the art. Such standard techniques can be found in reference manuals such as for example Sambrook et al. (1989, Molecular Cloning -

A Laboratory Manual, Cold Spring Harbor Laboratories) and Ausubel et al. (1994, Current Protocols in Molecular Biology, Wiley, New York).

The present description refers to a number of routinely used recombinant DNA (rDNA) technology terms. Nevertheless, definitions of selected examples of such rDNA terms are provided for clarity and consistency. For certainty, it is emphasized that the present invention finds utility with nucleic acids in general. Non-limiting examples of nucleic acids which can be used in accordance with the teachings of the present invention include that from eukaryotic cells such as that of animal cells, plant cells, or microorganisms as well as that from prokaryotic cells.

As used herein, the term "homopolymeric sequences" refers to a sequence composed essentially of a unique type of common base (adenosine A; cytosine C; guanine G; thymine T; uracil U) or of a less common base (non-limiting examples including inosine, I; and pseudouridine, Ψ).

As used herein, "nucleic acid molecule", refers to a polymer of nucleotides. Non-limiting examples thereof include DNA (i.e. genomic DNA, cDNA) and RNA molecules (i.e. mRNA). The nucleic acid molecule can be obtained by cloning techniques or synthesized. DNA can be double-stranded or single-stranded (coding strand or non-coding strand [antisense]).

The term "recombinant DNA" as known in the art refers to a DNA molecule resulting from the joining of DNA segments. This is often referred to as genetic engineering.

The terminology "amplification pair" refers herein to a pair of oligonucleotides (oligos) of the present invention, which are

sequence by one of a number of types of amplification processes, preferably a polymerase chain reaction. Other types of amplification processes include ligase chain reaction, strand displacement amplification, or nucleic acid sequence-based amplification, as explained in greater detail below. As commonly known in the art, the oligonucleotides are designed to bind to a complementary sequence under selected conditions.

The nucleic acid (i.e. DNA or RNA) for practicing the present invention may be obtained according to well known methods.

Oligonucleotide probes or primers of the present invention may be of any suitable length, depending on the particular assay format and the particular needs and targeted genomes employed. In general, the oligonucleotide probes or primers are at least 10 nucleotides in length, preferably between 12 and 24 molecules, and they may be adapted to be especially suited to a chosen nucleic acid amplification system. As commonly known in the art, the oligonucleotide probes and primers can be designed by taking into consideration the melting point of hydrizidation thereof with its targeted sequence (see below and in Sambrook et al., 1989, Molecular Cloning - A Laboratory Manual, 2nd Edition, CSH Laboratories; Ausubel et al., 1989, in Current Protocols in Molecular Biology, John Wiley & Sons Inc., N.Y.).

The term "oligonucleotide" or "DNA" molecule or sequence refers to a molecule comprised of the deoxyribonucleotides adenine (A), guanine (G), thymine (T) and/or cytosine (C), in a double-stranded form, and comprises or includes a "regulatory element" according to the present invention, as the term is defined herein. The term "oligonucleotide" or "DNA" can be found in linear DNA molecules or

fragments, viruses, plasmids, vectors, chromosomes or synthetically derived DNA. As used herein, particular double-stranded DNA sequences may be described according to the normal convention of giving only the sequence in the 5' to 3' direction.

As used herein, "oligonucleotides" or "oligos" define a molecule having two or more nucleotides (ribo or deoxyribonucleotides). In essence, "oligonucleotides" define at least dimers of nucleotides. The size of the oligonucleotide will be dictated by the particular situation and ultimately on the particular use thereof and adapted accordingly by the person of ordinary skill. An oligonucleotide can be synthesized chemically or derived by cloning according to well known methods.

Probes and oligonucleotides of the invention can be utilized with naturally occurring sugar-phosphate backbones as well as modified backbones including phosphorothioates, dithionates, alkyl phosphonates and α-nucleotides and the like. Modified sugar-phosphate backbones are generally taught by Miller, 1988, Ann. Reports Med. Chem. 23:295 and Moran et al., 1987, Nucleic acid molecule. Acids Res., 14:5019. Probes of the invention can be constructed of either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), and preferably of DNA. General teachings on the synthesis of oligonucleotides and substituents and modifications thereof can be found for example in US 5,438,131. The selection of the best suited synthesis pathway of an oligonucleotide and of the appropriate modifications, and substituents to be used, may be selected accordingly by the person of ordinary skill to which the instant invention pertains.

The modified oligonucleotides of the present invention can be synthesized chemically or produced through recombinant DNA

technology. All these methods are well known in the art. In a preferred embodiment, these modified oligonucleotides are a molecule composed of a single type of nucleotide (ribo- or deoxyribonucleotides, A, C, G, T or U) and containing at least one universal nucleoside. The length is dictated by the particular application. Typically, the oligonucleotide is at least 10 bases and probably 12-24 bases.

As used herein, a "primer" defines an oligonucleotide which is capable of annealing to a target sequence, thereby creating a double stranded region or duplex which can serve as an initiation point for DNA synthesis under suitable conditions.

"Nucleic acid hybridization" refers generally to the hybridization of two single-stranded nucleic acid molecules having complementary base sequences, which under appropriate conditions will form a thermodynamically favored double-stranded structure. Examples of hybridization conditions can be found in the two laboratory manuals referred above (Sambrook et al., 1989, *supra* and Ausubel et al., 1989, *supra*) and are commonly known in the art. In the case of a hybridization to a nitrocellulose filter, as for example in the well known Southern blotting procedure, a nitrocellulose filter can be incubated overnight at 65°C with a labeled probe in a solution containing 50% formamide, high salt (5 x SSC or 5 x SSPE), 5 x Denhardt's solution, 1% SDS, and 100 µg/ml denatured carrier DNA (i.e. salmon sperm DNA). The non-specifically binding probe can then be washed off the filter by several washes in 0.2 x SSC/0.1% SDS at a temperature which is selected in view of the desired stringency: room temperature (low stringency), 42°C (moderate stringency) or 65°C (high stringency). The selected temperature is based on the melting temperature (Tm) of the DNA hybrid.

Of course, RNA-DNA hybrids can also be formed and detected. In such cases, the conditions of hybridization and washing can be adapted according to well known methods by the person of ordinary skill. Stringent conditions will be preferably used (Sambrook et al.,1989, *supra*).

The types of detection methods in which probes can be used include Southern blots (DNA detection), dot or slot blots (DNA, RNA), and Northern blots (RNA detection).

Although the present invention is not specifically dependent on the use of a label, such a label might be beneficial in certain embodiments. Probes or oligonucleotides can be labeled according to numerous well known methods (Sambrook et al., 1989, supra). Non-limiting examples of labels include ^{3}H, ^{14}C, ^{32}P, and ^{35}S. Non-limiting examples of detectable markers include ligands, fluorophores, chemiluminescent agents, enzymes, and antibodies. Other detectable markers for use with probes, which can enable an increase in sensitivity of the method of the invention, include biotin and radionucleotides. It will become evident to the person of ordinary skill that the choice of a particular label dictates the manner in which it is bound to the probe.

As commonly known, radioactive nucleotides can be incorporated into probes of the invention by several methods. Non-limiting examples thereof include kinasing the 5' ends of the probes using gamma ^{32}P ATP and polynucleotide kinase, using the Klenow fragment of Pol I of *E. coli* in the presence of radioactive dNTP (i.e. uniformly labeled DNA probe using random oligonucleotide primers in low-melt gels), using the SP6/T7 system to transcribe a DNA segment in the presence of one or more radioactive NTP, and the like.

Amplification of a selected, or target, nucleic acid sequence may be carried out by a number of suitable methods. See generally Kwoh et al., 1990, Am. Biotechnol. Lab. 8:14-25. Numerous amplification techniques have been described and can be readily adapted to suit particular needs of a person of ordinary skill. Non-limiting examples of amplification techniques include polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA), transcription-based amplification, the Qβ replicase system and NASBA (Kwoh et al., 1989, Proc. Natl. Acad. Sci. USA 86, 1173-1177; Lizardi et al., 1988, BioTechnology 6:1197-1202; Malek et al., 1994, Methods Mol. Biol., 28:253-260; and Sambrook et al., 1989, *supra*). Preferably, amplification will be carried out using PCR.

Polymerase chain reaction (PCR) is carried out in accordance with known techniques. See, e.g., U.S. Pat. Nos. 4,683,195; 4,683,202; 4,800,159; and 4,965,188 (the disclosures of all three U.S. Patent are incorporated herein by reference). In general, PCR involves, a treatment of a nucleic acid sample (e.g., in the presence of a heat stable DNA polymerase) under hybridizing conditions, with one oligonucleotide primer for each strand of the specific sequence to be detected. An extension product of each primer which is synthesized is complementary to each of the two nucleic acid strands, with the primers sufficiently complementary to each strand of the specific sequence to hybridize therewith. The extension product synthesized from each primer can also serve as a template for further synthesis of extension products using the same primers. Following a sufficient number of rounds of synthesis of extension products, the sample is analysed to assess whether the sequence or sequences to be detected are present.

Detection of the amplified sequence may be carried out by visualization following EtBr staining of the DNA following gel electrophores, or using a detectable label in accordance with known techniques, and the like. For a review on PCR techniques (see PCR Protocols, A Guide to Methods and Amplifications, Michael et al. Eds, Acad. Press, 1990).

Ligase chain reaction (LCR) is carried out in accordance with known techniques (Weiss, 1991, Science 254:1292). Adaptation of the protocol to meet the desired needs can be carried out by a person of ordinary skill. Strand displacement amplification (SDA) is also carried out in accordance with known techniques or adaptations thereof to meet the particular needs (Walker et al., 1992, Proc. Natl. Acad. Sci. USA 89:392-396; and ibid., 1992, Nucleic Acids Res. 20:1691-1696).

As used herein, the term "gene" is well known in the art and relates to a nucleic acid sequence defining a single protein or polypeptide. A "structural gene" defines a DNA sequence which is transcribed into RNA and translated into a protein having a specific amino acid sequence thereby giving rise the a specific polypeptide or protein. It will be readily recognized by the person of ordinary skill, that the nucleic acid sequence of the present invention can be incorporated into anyone of numerous established kit formats which are well known in the art.

The term "vector" is commonly known in the art and defines a plasmid DNA, phage DNA, viral DNA and the like, which can serve as a DNA vehicle into which DNA of the present invention can be cloned. Numerous types of vectors exist and are well known in the art.

The term "allele" defines an alternative form of a gene which occupies a given locus on a chromosome.

As commonly known, a "mutation" is a detectable change in the genetic material which can be transmitted to a daughter cell. As well known, a mutation can be, for example, a detectable change in one or more deoxyribonucleotide. For example, nucleotides can be added, deleted, substituted for, inverted, or transposed to a new position. Spontaneous mutations and experimentally induced mutations exist. The result of a mutations of nucleic acid molecule is a mutant nucleic acid molecule. A mutant polypeptide can be encoded from this mutant nucleic acid molecule.

As used herein, the term "purified" refers to a molecule having been separated from a cellular component. Thus, for example, a "purified protein" has been purified to a level not found in nature. A "substantially pure" molecule is a molecule that is lacking in all other cellular components.

The present invention also relates to a kit comprising the oligonucleotide primers of the present invention. For example, a compartmentalized kit in accordance with the present invention includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample (DNA or cells), a container which contains the primers used in the assay, containers which contain enzymes, containers

which contain wash reagents, and containers which contain the reagents used to detect the extension products.

BRIEF DESCRIPTION OF THE DRAWINGS

Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which:

Figure 1 (PRIOR ART) shows an example of the steps involved in generating cDNA libraries from mRNA. Although a number of strategies can be used for cDNA library generation, of which only one is shown above, all libraries require as a first step, a primer from which the reverse transcriptase (RT) can prime. In the case of full-length cDNA libraries, an oligo d(T) primer is used because it anneals to the 3' poly (A) tail of the eukaryotic mRNAs. In the case of prokaryotic, some viral, or other eukaryotic mRNAs which lack a poly (A) tail, a homopolymeric stretch of nucleoside 5'-monophosphates can be added to the 3' end of the mRNA. For example, poly (A) polymerase can be used to add a poly (A) tail to mRNAs which lack one. An oligonucleotide which contains complementary nucleotides (e.g. oligo d(T)) is then annealed to the mRNA and serves as primer for the RT.

Figure 2A shows hybridization of oligo $d(T)_{15}$ primer to the *bona fide* poly (A) tail of an mRNA (right) or to an internal A-rich stretch (left) within the mRNA by conventional oligo d(T) primer used in current cDNA library construction. Although the length of the primer used can differ, and the two 3' most nucleotides are sometimes (A,C,G,T) and (A,G,C) to "lock" the oligonucleotide in place at the junction of the body of the mRNA and the poly (A) tail, neither of these modifications prevent

the misannealing of the oligo d(T) primer to internal A-rich stretches. The asterisks denotes mispairing resulting in destabilization of the duplex. Figure 2B shows the chemical structure of 3-nitropyrrole. Figure 2C shows the structure of oligo d(T)•Z primer. Figure 2D shows the expected discrimination between the poly (A) tail (right) and internal A-rich stretches (left) when hybridizing to oligo d(T)•Z. The asterisks denote mispairing resulting in destabilization of the duplex and circles represent 3-nitropyrrole artificial mismatches.

Figure 3A shows the structure of the eIF-4GII cDNA construct used to analyze mispriming at the 3' end. The location of four internal A-rich sequences are shown - all of which generated 3' truncated clones when eIF-4GII was isolated from a cDNA library. The plasmid was linearized with Asp 718 and T7 RNA polymerase used to generate a ~2400 nt ^{3}H-test transcript. Figure 3B shows the integrity of the *in vitro* generated transcript following fractionation on a formaldehyde/ 1.2% agarose gel, treatment with EN^{3}HANCE, and autoradiography of the dried gel. Figure 3C shows the alkaline agarose analysis of RT products generated by priming synthesis with oligo d(T) (lane 1) or oligo d(T)•Z (lane 2) using MMLV RT. Complementary DNA was labeled with α-^{32}P-dCTP. The position of migration of truncated products are indicated by a filled circle and full length product by an arrow. These results directly demonstrate correction of 3' mispriming by utilizing oligo d(T)•Z as primer during first strand synthesis.

Figure 4A shows the structure of eIF-4GII construct used to demonstrate mispriming at the 3' end. The location of five oligonucleotides (a, b, c, d, e) used in the hybridization assay to map the sites of 3' mispriming by oligo d(T) are shown. The nucleotide targets of

the oligonucleotides on eIF-4GII are: Oligo a, ^{5567}GAAATTGACTCAGTACTATT^{5584}; Oligo b, ^{5416}GAAGGAAATGCTGTGGACC^{5535}; Oligo c, ^{5194}TGTATAATAGAAAAGCAGAG^{5213}; Oligo d, ^{5068}TTTTAAACAAGGACTCATAC^{5087}; and Oligo e, ^{4781}AAGAGGAGTCTGAGGATAAC^{4800}. Figure 4B shows the Southern blot of the alkaline agarose gel of RT products generated by priming synthesis with either oligo d(T) or oligo d(T)•Z. Marker lane refers to the 1 kb size ladder from GIBCO and sizes (in bp) are indicated to the left of the diagram. eIF-4GII DNA refers to a DNA fragment of eIF-4GII used as a positive control for DNA hybridizations. Oligonucleotides used as probes on each blot are indicated below each panel. The asterisks on the left denotes the cDNA product obtained by priming at the correct poly (A) site. The filled circle denotes the cDNA product obtained by priming from the A-rich stretch between nucleotides 5550-5575, whereas the arrow denoted the cDNA obtained by priming from nucleotides 5085-5120.

Figure 5 shows an example illustrating mispriming events at the 5' end of cDNAs during cDNA library construction of 5' RACE analysis to extend the sequence of known genes.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawing which is exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The demonstration of the destabilization effect of non-specific or artifactual duplex formation and of its concurrent effect on mismatch and/or mispriming events was carried out with an oligo d(T) primer, modified with one or two universal analogues. Whether such an introduction could result in increased discrimination between the perfectly matched target of that primer (i.e. the 3' poly (A) tail of the mRNA) and an imperfect matched sequenced (internal A-rich stretches) was analyzed.

More specifically, an oligo d(T) primer, called oligo d(T)•Z was generated, in which two of the thymine bases were substituted by 3-nitropyrrole (Fig. 2C). General teachings on 3-nitropyrrole, the synthesis thereof and the like can be found for example in U.S. 5,438,131.

To test whether this primer can reduce mispriming from internal A-rich sequences (Fig. 2D), a cDNA clone from eIF-4GII, a eukaryotic translation factor, was obtained. When cDNA clones to this gene were initially isolated, only one of 5 clones had the correct 3' end. Sequence characterization of these clones demonstrated that all the truncated clones were the result of internal priming by oligo d(T) at four different sites (denoted as leftward arrows in Fig. 3A). *In vitro* transcribed RNA generated from this clone thus serves as an excellent test reagent to determine the ability of the 3-nitropyrrole substituted oligo d(T) to decrease the number of mispriming events. The quality of the *in vitro* transcribed RNA is shown in Fig. 3B and demonstrates that the test template is intact. This RNA was then annealed to oligo d(T) or oligo d(T)•Z, and reverse transcription performed with MMLV RT. As shown in Fig. 3C, use of oligo d(T) on this template resulted in shorter than full-

length products (>95%) generated as a result of internal priming (Fig. 3C, lane 1). However, use of oligo d(T)•Z as primer on the same template resulted in the majority (>95%) of products being full-length (Fig. 3C, lane 2).

These results demonstrate that use of oligo d(T)•Z in reverse transcription reactions significantly improves the specificity for the 3' poly (A) tail and demonstrates the usefulness of this procedure in destabilizing non-specific duplex formation and more particularly for generating full length cDNAs.

The sites of mispriming with oligo d(T) on the control eIF-4GII template were identified (Fig. 4). This was done by fractionating the products of RT reactions performed with either oligo d(T) or oligo d(T)•Z on an alkaline agarose gel followed by transfer to a nylon membrane. This membrane was then probed, by hybridization, with oligonucleotides designed to target various regions of the 3' untranslated region of eIF-4GII (oligonucleotides are labelled a, b, c, d and e in Fig. 4A). As shown in Fig. 4B, hybridization with oligonucleotide "a" detected correctly primed cDNA when both oligo d(T) and oligo d(T)•Z were used as primer. Hybridization with oligonucleotides b and c, detected a novel truncated product when the RT reaction was primed with oligo d(T), indicating mispriming from an internal A-rich stretch with this primer (Fig. 4B). Hybridization with oligonucleotides d and e, detected an additional novel, more abundant truncated product (denoted by arrowheads in Fig. 4B) when the RT reaction was primed with oligo d(T), indicating mispriming from a second internal A-rich stretch with this primer but not with oligo d(T)•Z (Fig. 4B).

Mispriming event are common in Rapid Amplification of cDNA ends (RACE). An example of mispriming at the 5' end of cDNAs during 5' RACE analysis is shown in Figure 5. Such mispriming events could be resolved by incorporating a universal nucleoside into the oligo d(C) primer to increase the discrimination between the homologous target (i.e. - the 5' end G tail) and an internal G-rich sequence. It is expected that incorporation of at least one universal base (i.e. 3-nitropyrrole) in the homopolymeric oligo d(C) primer should significantly reduce such mispriming.

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

WHAT IS CLAIMED IS:

1. A method for destabilizing non-specific duplex formation between an oligonucleotide and a target nucleic acid, wherein at least one of said oligonucleotide and target nucleic acid is a homopolymeric sequence comprising an incubation of said target nucleic acid with a modified oligonucleotide, wherein said modified oligonucleotide includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

2. The method of claim 1, wherein said modification is at least one universal base incorporated into said oligonucleotide.

3. The method of claim 2, wherein said universal base is 3-nitropyrrole.

4. The method of one of claims 1-3, wherein said oligonucleotide is a homopolymer.

5. A method for increasing the proportion of full length cDNA clones in a library, comprising a use of a modified oligo d(T) during first strand synthesis, wherein said modification decreases or abrogates hydrogen bonding between said modified oligo d(T) and a non-specific target sequence, thereby increasing the proportion of full length cDNA clones.

6. The method of claim 5, wherein said modification is at least one universal base incorporated into said oligo d(T).

7. The method of claim 6, wherein said universal base is 3-nitropyrrole.

8. The method of claim 5, wherein said modification is at least one chemically modified nucleoside incorporated into said oligo d(T).

9. The method of claim 5, wherein said modification is at least one base analog (e.g. inosine) incorporated into said oligo d(T).

10. The method of claim 8, wherein said base analog is inosine.

11. The method of claim 5, wherein said modification is at least one mismatch incorporated into said oligo d(T).

12. The method of claim 5, wherein said modification is a phosphate or ribose modification incorporated into said oligo d(T).

13. The method according to claim 5, wherein an enzyme capable of RNA-dependent DNA polymerization is used for said first strand synthesis.

14. The method according to claim 13, wherein said enzyme is a reverse transcriptase selected from the group consisting of avian myoblastoid virus reverse transcriptase, murine moloney leukemia virus reverse transcriptase, and human immuno deficiency virus reverse transcriptase.

15. A kit for the synthesis of cDNA, said kit comprising a modified oligo d(T) primer, wherein said modified oligonucleotide includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

16. A method for reducing mispriming events during DNA synthesis comprising a use of a modified oligonucleotide to prime said DNA synthesis, wherein said modification decreases or abrogates hydrogen bonding between said modified oligonucleotide and a non-specific target sequence, thereby reducing mispriming events, and wherein at least one of said modified oligonucleotide and specific target sequence is a homopolymer.

17. The method of claim 16, wherein said modification is at least one universal base incorporated into said oligonucleotide.

18. The method of claim 17, wherein said universal base is 3-nitropyrrole.

19. The method of claims 17, 18 or 19, wherein said oligonucleotide is a homopolymer.

20. A method for reducing mispriming during 5' RACE comprising a use of a modified oligonucleotide to prime said 5' RACE, wherein said modification decreases or abrogates hydrogen bonding between said modified oligonucleotide and a non-specific target sequence, thereby reducing mispriming events, and wherein at least of said modified oligonucleotide and specific target sequence is a homopolymer.

21. The method of claim 20, wherein said modification is at least one universal base incorporated into said oligonucleotide.

22. The method of claim 20, wherein said universal base is 3-nitropyrrole.

23. The method of claim 20, wherein said modification is at least one chemically modified nucleoside incorporated into said oligonucleotide.

24. The method of claim 20, wherein said modification is at least one base analog incorporated into said oligonucleotide.

25. The method of claim 24, wherein said base analog is inosine.

26. The method of claim 20, wherein said modification is at least one mismatch incorporated into said oligonucleotide.

27. The method of claim 20, wherein said modification is a phosphate or ribose modification destabilizing mismatch recognition incorporated into said oligonucleotide.

28. A kit for 5' RACE comprising a modified homopolymeric primer, wherein said modified homopolymeric primer includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

29. A method for reducing mispriming during 3' RACE comprising a use of a modified oligonucleotide to prime said 3' RACE, wherein said modification decreases or abrogates hydrogen bonding between said modified oligonucleotide and a non-specific target sequence, thereby reducing mispriming events, and wherein at least of said modified oligonucleotide and specific target sequence is a homopolymer.

30. The method of claim 29, wherein said modification is at least one universal base incorporated into said oligonucleotide.

31. A method for generating genetic markers comprising a use of a modified oligonucleotide to prime from homopolymeric

stretches, thereby decreasing or abrogating hydrogen bonding between said modified oligonucleotide and said homopolymeric stretches.

32. The method of claim 31, wherein said modified oligonucleotide primes from an internal A-rich region in an Alu repeat.

33. A method for stabilizing duplex formation between an oligonucleotide and a target homopolymeric sequence comprising an incubation of said target homopolymeric sequence with a modified oligonucleotide, wherein said modified oligonucleotide includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

34. A method for reducing mispriming during sequencing comprising a use of a modified oligonucleotide to prime DNA synthesis, wherein said modified oligonucleotide includes modifications which decrease or abrogate hydrogen bonding between same and non-specific target sequences.

cDNA Library Generation

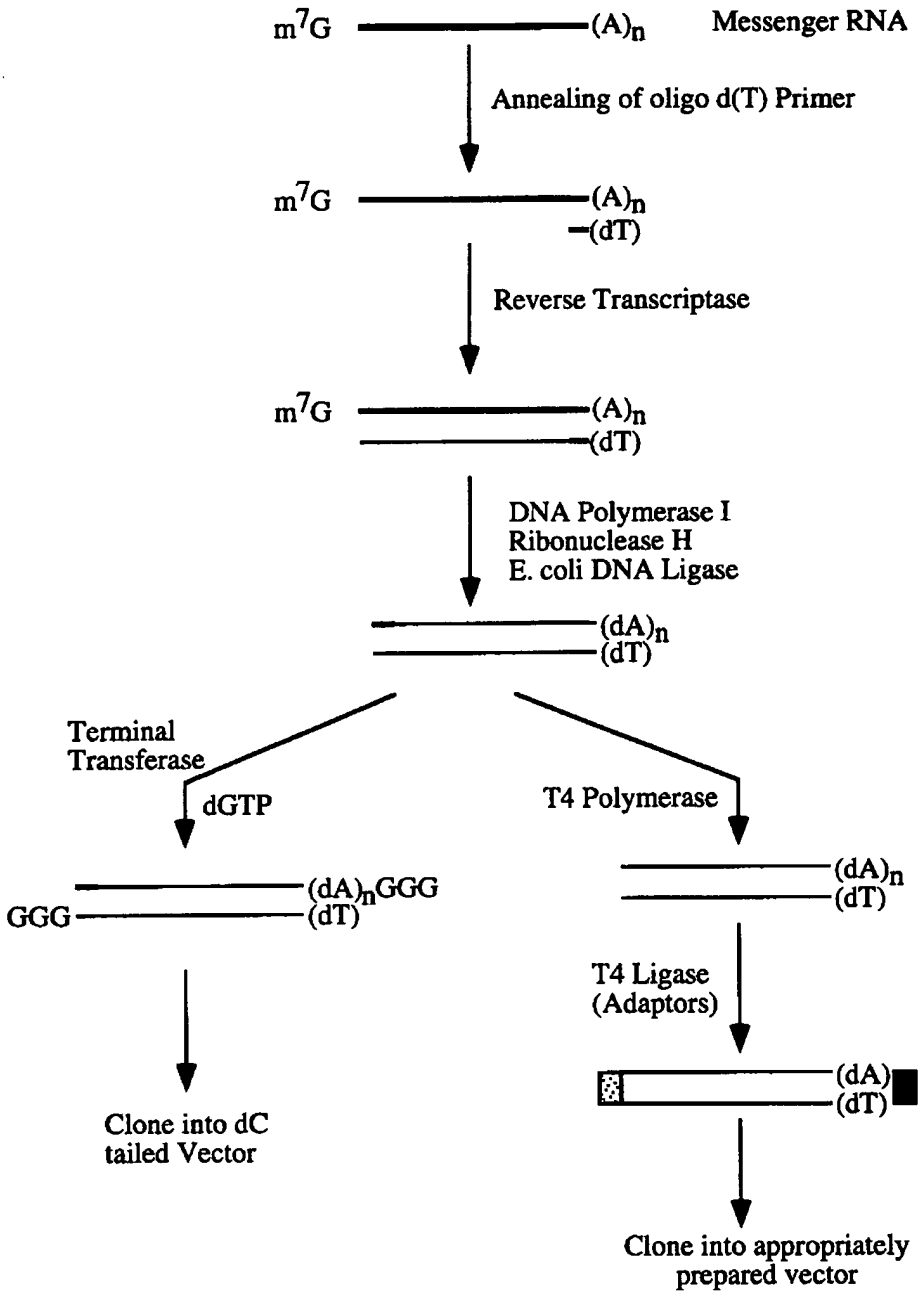
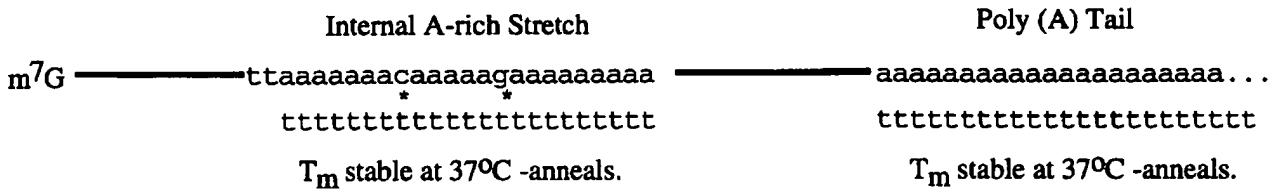


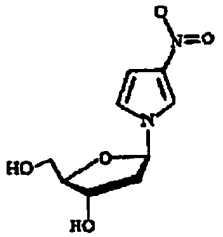
Fig. 1. An Example of the Steps Involved in Generating cDNA Libraries from mRNA. Although a number of strategies can be used for cDNA library generation, of which only one is shown above, all libraries require as a first step, a primer which Reverse Transcriptase (RT) can prime from. In the case of full-length cDNA libraries, an oligo d(T) primer is used because it anneals to the 3' poly (A) tail of the eukaryotic mRNAs. In the case of prokaryotic, some viral, or other eukaryotic mRNAs which lack a poly (A) tail, a homopolymeric stretch of nucleosides can be added to the 3' end of the mRNA. For example, poly (A) polymerase can be used to add a poly (A) tail to mRNAs which lack one. Use of an oligonucleotide which contains complementary nucleotides (e.g. oligo d(T)) is then annealed to the mRNA and serves as primer for the RT.

Design and hybridization of a modified oligo d(T) primer containing a universal nucleoside

A.



B.



Z = 1-(2'-Deoxy-β-D-Ribofuranosyl)-3-Nitropyrrole

C. Oligo d(T)•Z Primer: $^{5'}$ d(T)$_7$•Z•d(T)$_9$•Z•d(T)$_5$ $^{3'}$

D.

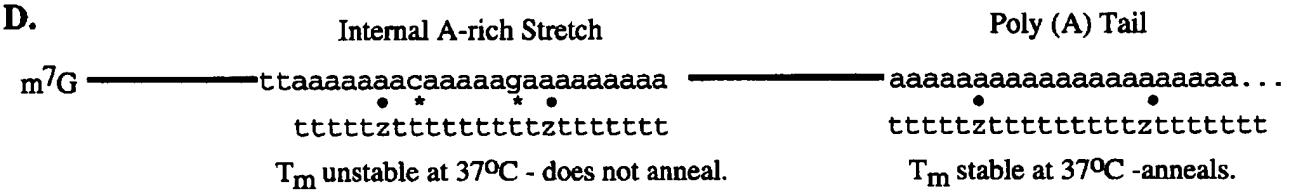


Fig. 2. A) Hybridization of oligo d(T)$_{15}$ primer to the *bona fide* poly (A) tail of an mRNA (right) or to an internal A-rich stretch (left) within the mRNA. Conventional oligo d(T) primer used in current cDNA library construction. Although the length of the primer used can differ, and the two 3' most nucleotides are sometimes (A,C,G,T) and (A,G,C) to "lock" the oligo in place at the junction of the body of the mRNA and the poly (A) tail, neither of these modifications alter the misannealing of the oligo d(T) primer to internal A-rich stretches. The asterisks denotes mispairing resulting in destabilization of the duplex. **B)** Chemical structure of 3-nitropyrrole **C)** Structure of oligo d(T)•Z primer. **D)** Expected discrimination expected between the poly (A) tail (right) and internal A-rich stretches (left) when hybridizing to oligo d(T)•Z. The asterisks denotes mispairing resulting in destabilization of the duplex and circles represent 3-nitropyrrole artificial mismatches.

Correction of 3' mispriming by using a 3-nitropyrrole substituted oligo d(T) primer

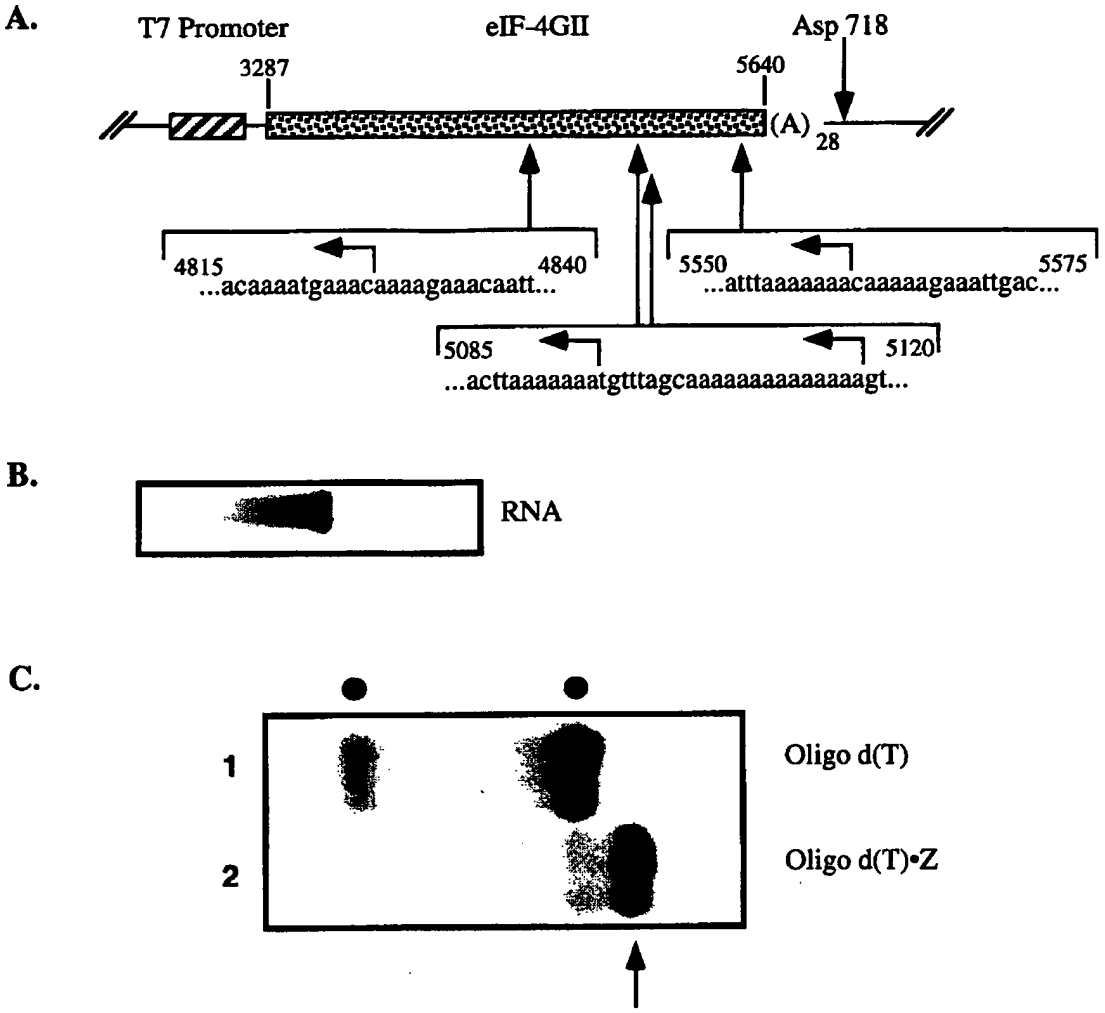


Fig. 3. A) Structure of eIF-4GII construct used to analyze mispriming at the 3' end. The location of four internal A-rich sequences are shown - all of which generated 3' truncated clones when eIF-4GII was isolated from a cDNA library. The plasmid was linearized with Asp 718 and T7 RNA polymerase used to generate a ~2400 nt ^{3}H-test transcript. **B)** The integrity of the *in vitro* generated transcript is shown following fractionation on a formaldehyde/ 1.2% agarose gel, treatment with EN^{3}HANCE, and autoradiography of the dried gel. **C)** Alkaline agarose analysis of RT products generated by priming synthesis with oligo d(T) (lane 1) or oligo d(T)•Z (lane 2) using MMLV Superscript. Complementary DNA was labeled with α-^{32}P-dCTP. The position of migration of truncated products are indicated by a filled circle and full length product by an arrow.

Mapping of sites of mispriming by oligo d(T) on eIF-4GII test transcript

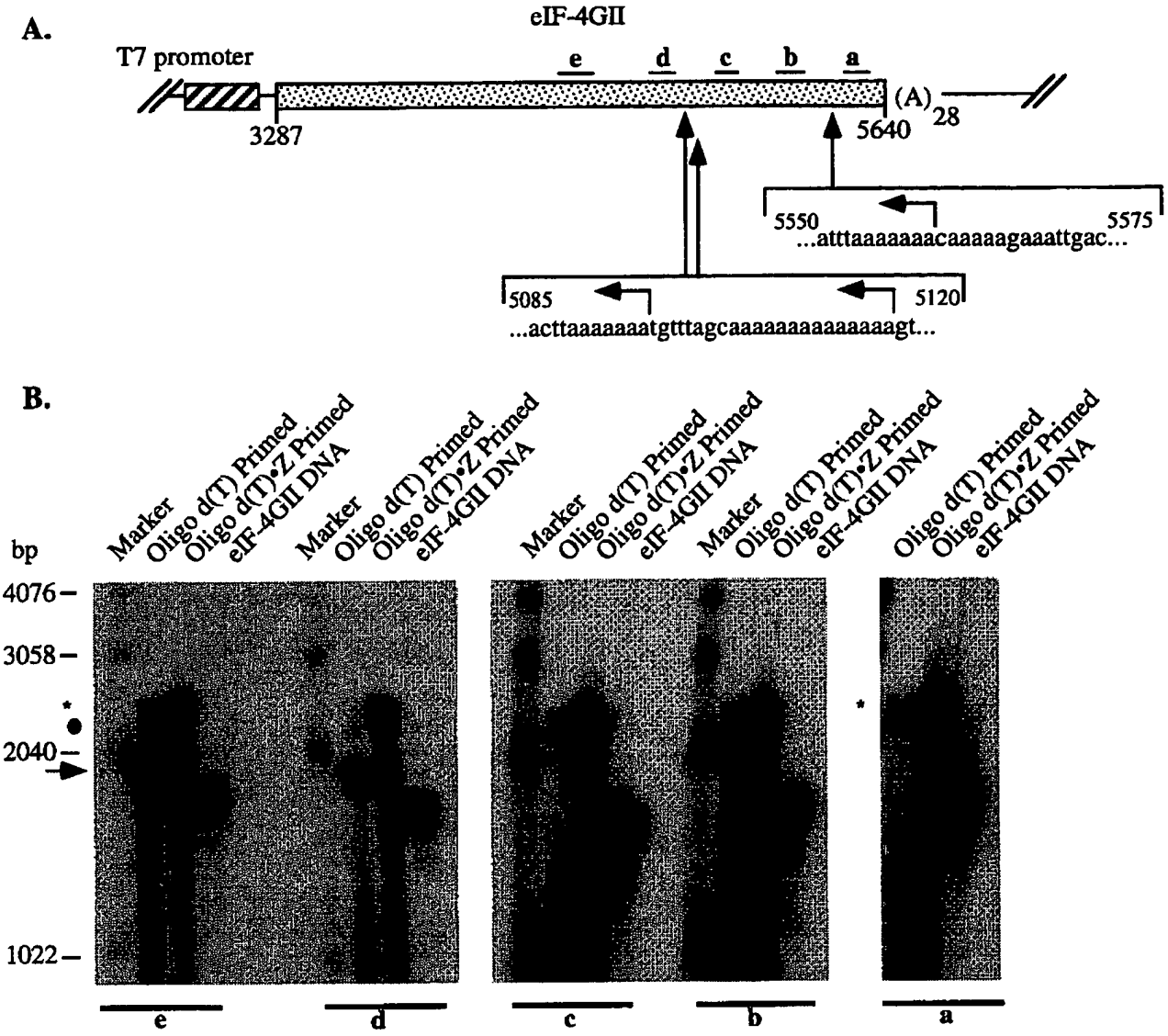


Fig. 4. A) Structure of eIF-4GII construct used to demonstrate mispriming at the 3' end. The location of five oligos (a, b, c, d, e) used in the hybridization assay to map the sites of 3' mispriming by oligo d(T) are shown. The nucleotide targets of the oligonucleotides on eIF-4GII are: Oligo a, ^{5567}GAAATTGACTCAGTACTATT$_{5584}$; Oligo b, ^{5416}GAAGGAAATGCTGTGGACC$_{5535}$; Oligo c, ^{5194}TGTATAATAGAAAAGCAGAG$_{5213}$; Oligo d, ^{5068}TTTTAAACAAGGACTCATAC$_{5087}$; and Oligo e, ^{4781}AAGAGGAGTCTGAGGATAAC$_{4800}$. **B)** Southern blot of Alkaline agarose gel of RT products generated by priming synthesis with either oligo d(T) or oligo d(T)•Z. Marker lane refers to the 1 kb size ladder from GIBCO and sizes (in bp) are indicated to the left of the diagram. eIF-4GII DNA refers to a DNA fragment of eIF-4GII used as a positive control for DNA hybridizations. Oligos used as probes on each blot are indicated below each panel. The asterisks on the left denotes the cDNA product obtained by priming at the correct poly (A) site. The filled circle denotes the cDNA product obtained by priming from the A-rich stretch between nucleotides 5550-5575, whereas the arrow denoted the cDNA obtained by priming from nucleotides 5085-5120.

An example of a mispriming event at the 5' end during cDNA synthesis for the purposes of cloning or RACE analysis

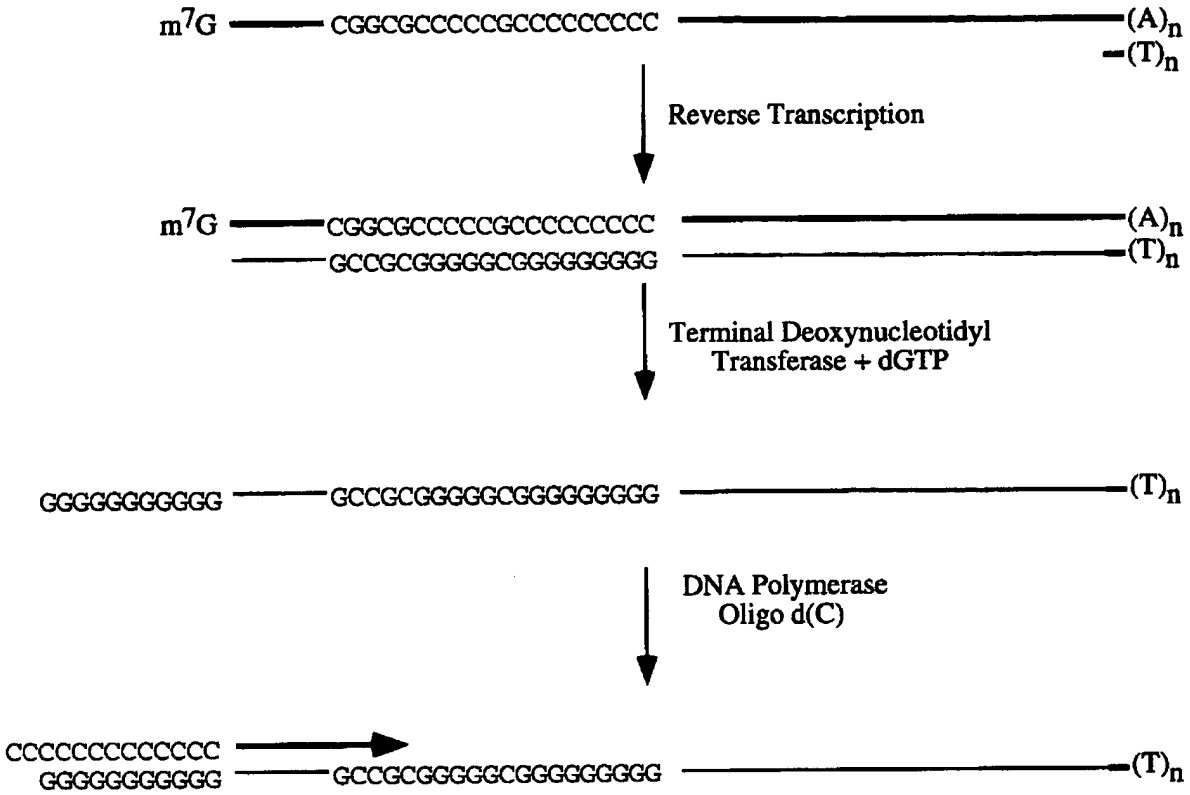


Fig. 5. An example illustrating mispriming events at the 5' end of cDNAs during cDNA library construction of 5' RACE analysis to extend the sequence of known genes. Such mispriming events could be resolved by incorporating a universal nucleoside into the oligo d(C) primer to increase the discriminationbetween the homologous target (i.e. - the 5' end G tail) and an internal G-rich sequence. This idea also pertains to the use of any homopolymeric primer targeting its complement.